Atmospheric Science Data Center Update

CERES Science Team Meeting

Jeff Walter

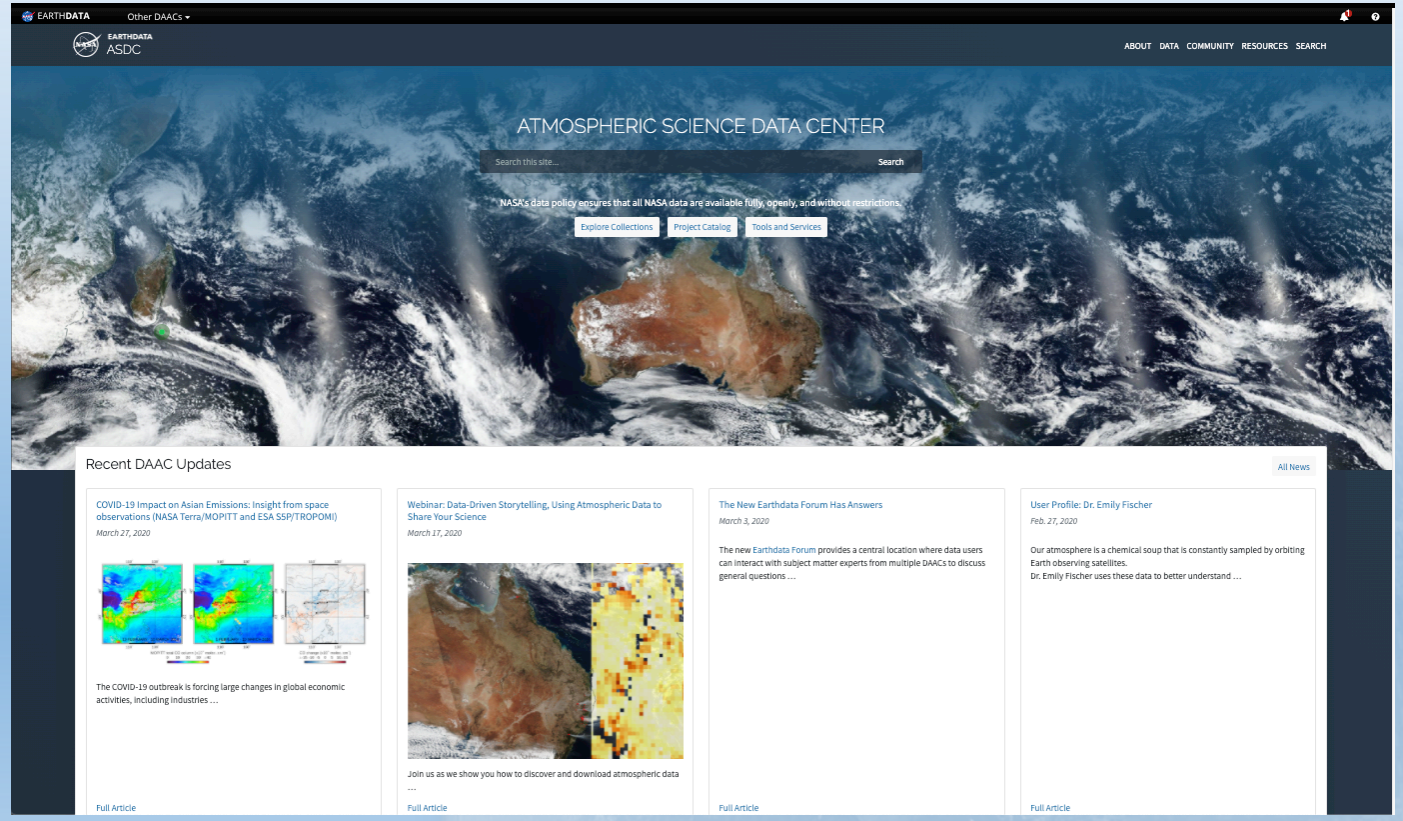
04/28/2020

Outline

- New Capabilities
- Infrastructure Updates
- ASDC in the Cloud

New ASDC Web Site

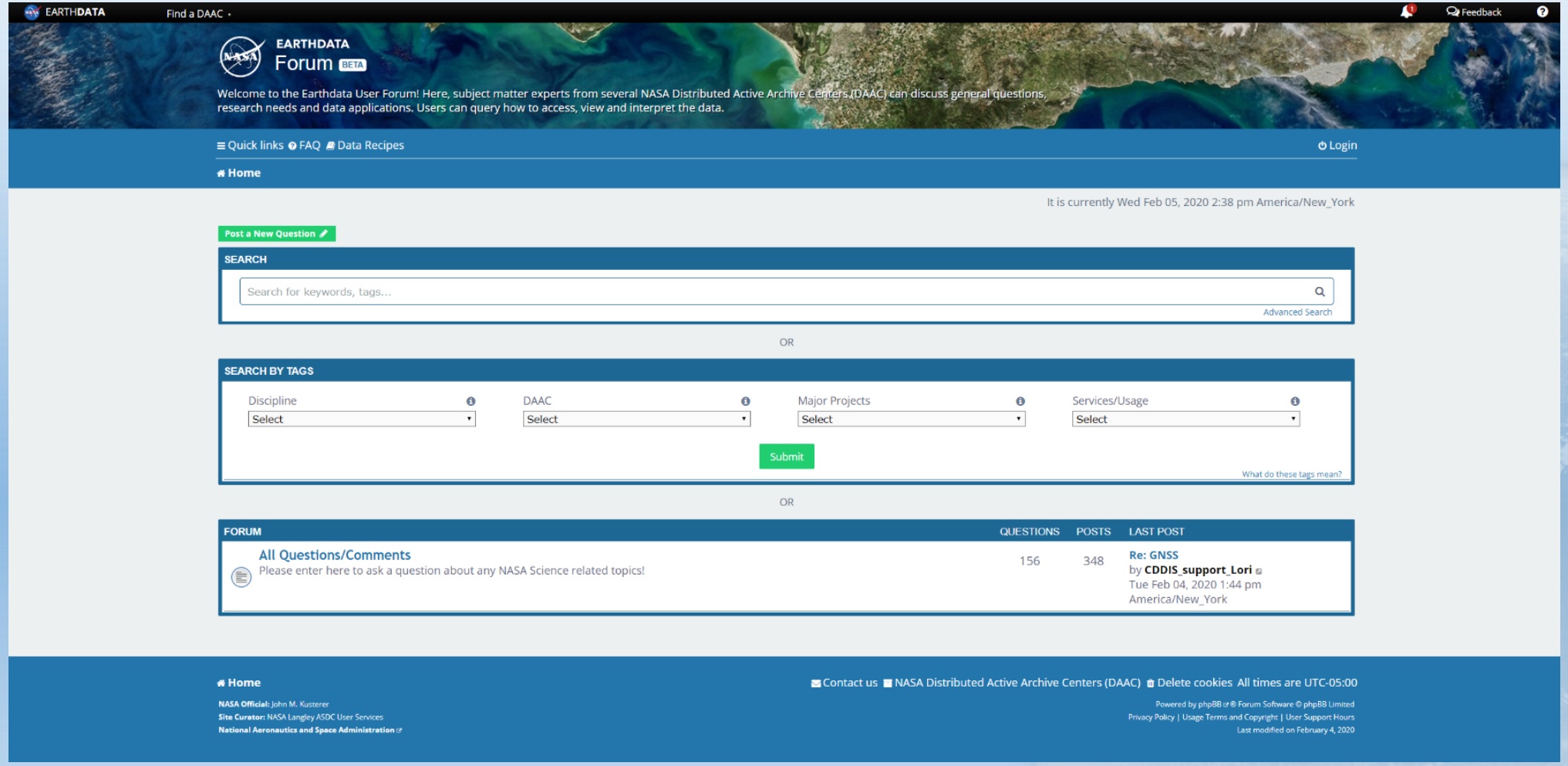
- New ASDC website:
  - Currently in Beta version
- https://asdc.larc.nasa.gov
- Each mission has a Dynamic Landing Page (DLP)
- Tools & Services page → Learning Resources
  - Microarticles, Jupyter Notebooks, StoryMaps, etc.



Earthdata Forum

- https://forum.earthdata.nasa.gov/
- Users post questions, moderators assign tags, SMEs answer technical details related to data
- Login with Earthdata login
- Can search by keyword; Drop down menus for Discipline, DAAC, Major Projects, and Services/Usage
- Launched on February 25th

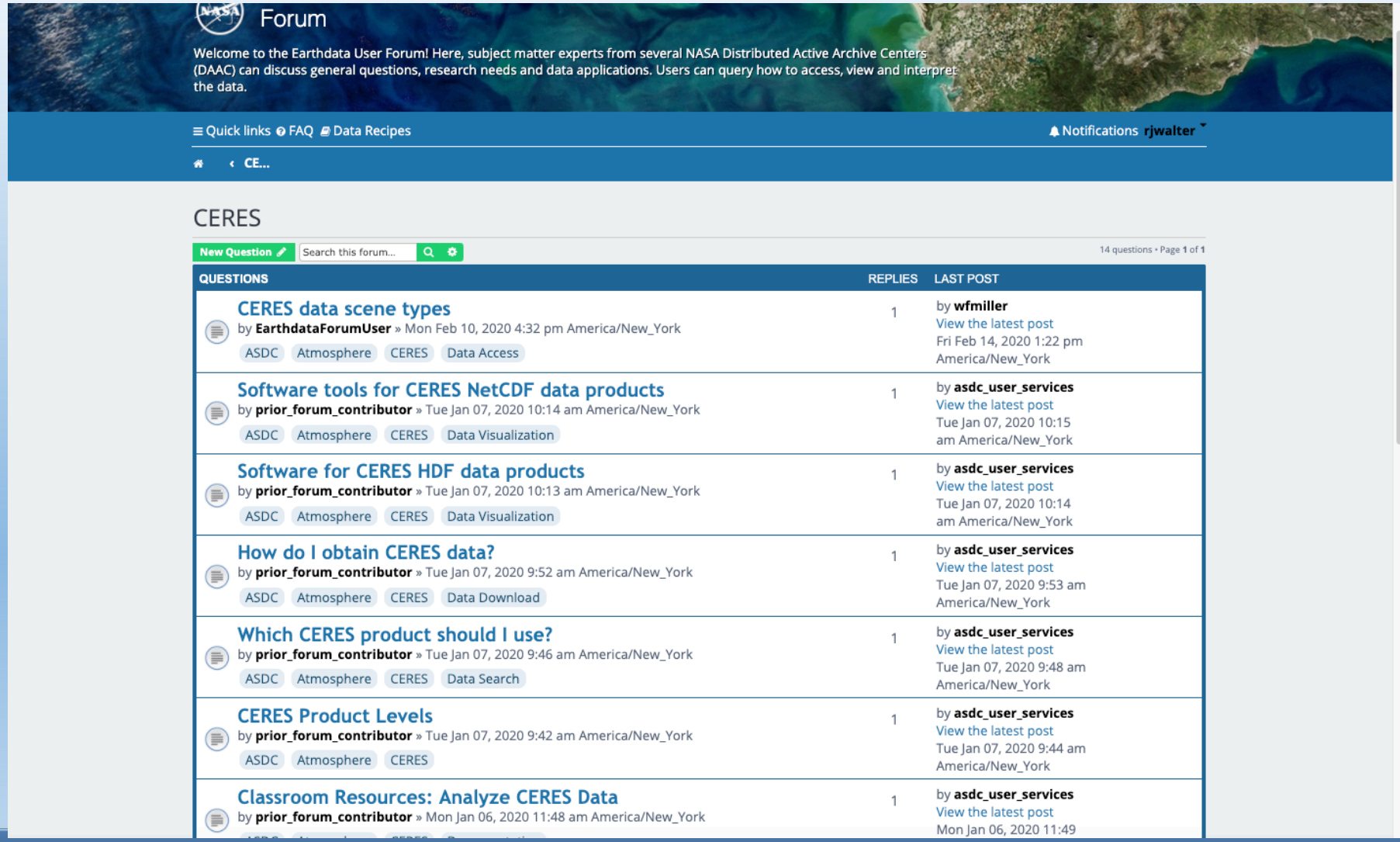
Earthdata Forum



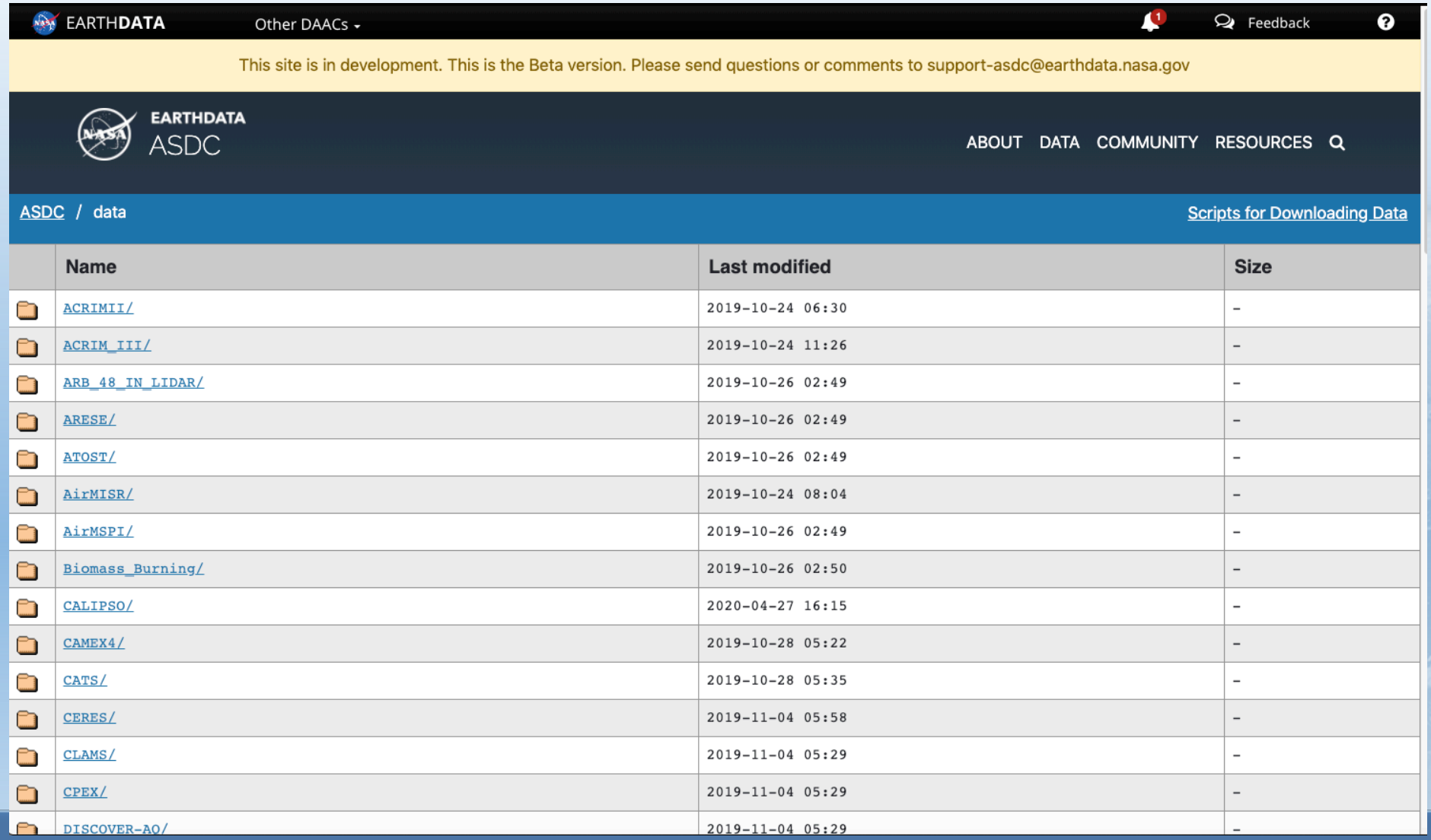
Atmospheric Science Data Center

NASA Langley Research Center
Hampton, VA

Earthdata Forum

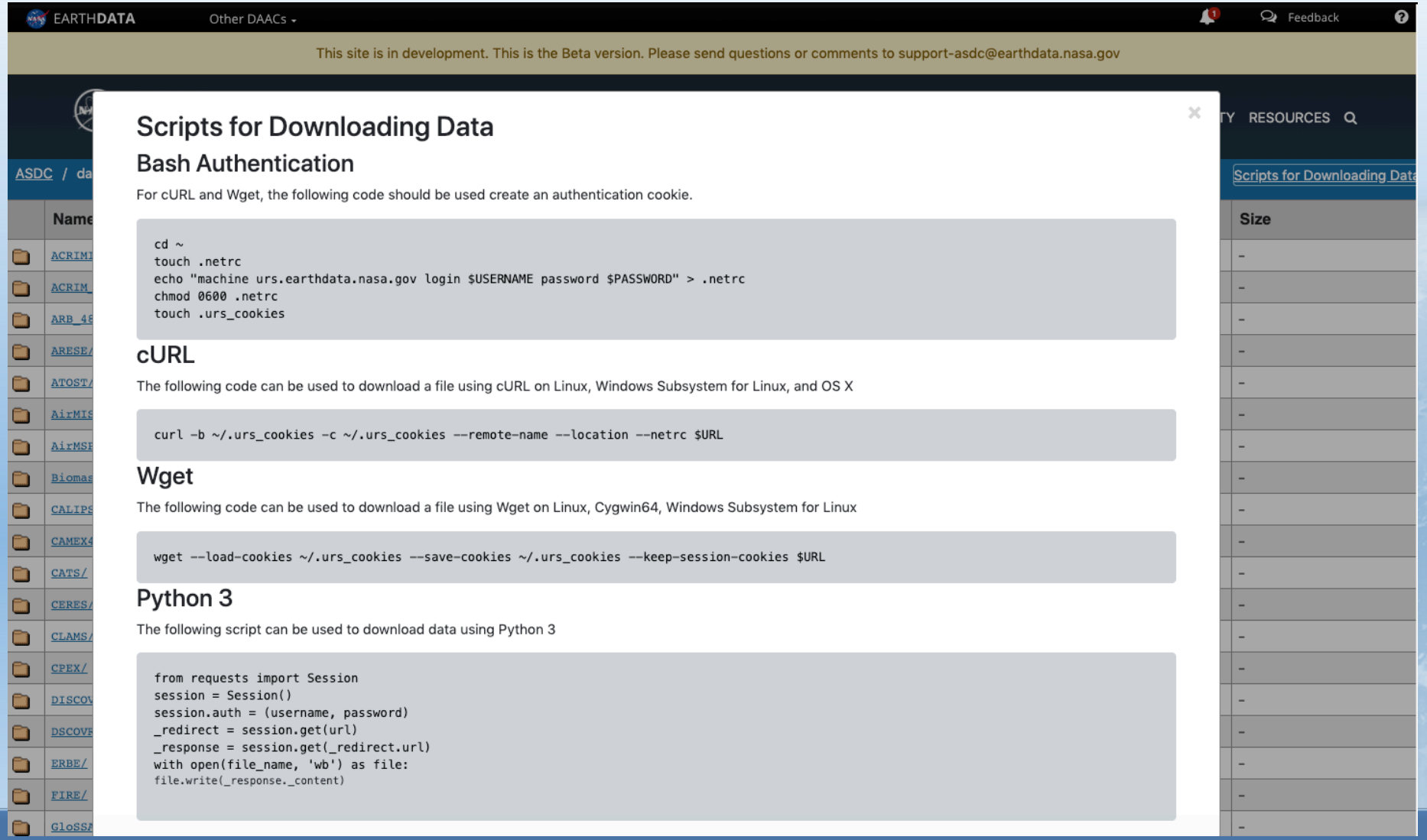


Direct Data Download



	Name	Last modified	Size
	ACRIMII/	2019-10-24 06:30	-
	ACRIM_III/	2019-10-24 11:26	-
	ARB_48_IN_LIDAR/	2019-10-26 02:49	-
	ARESE/	2019-10-26 02:49	-
	ATOST/	2019-10-26 02:49	-
	AirMISR/	2019-10-24 08:04	-
	AirMSPI/	2019-10-26 02:49	-
	Biomass_Burning/	2019-10-26 02:50	-
	CALIPSO/	2020-04-27 16:15	-
	CAMEX4/	2019-10-28 05:22	-
	CATS/	2019-10-28 05:35	-
	CERES/	2019-11-04 05:58	-
	CLAMS/	2019-11-04 05:29	-
	CPEX/	2019-11-04 05:29	-
	DISCOVER-AQ/	2019-11-04 05:29	-

Direct Data Download



Infrastructure Updates

- Retiring the WebRA environment and moving web capabilities to either OpenShift or the VM stack
  - OPeNDAP and ArcGIS moving to VMs
  - POWER moving to VM first and eventually to OpenShift
- 6 PB storage in the DMZ
  - Making all public data available via Direct Data Download (https) in the DMZ
- Future plans
  - Virtualized infrastructure
  - Move all (or nearly all) infrastructure to the DMZ

Infrastructure Updates

- OpenShift
  - Moved underlying application nodes to VMs
    - Allows for live migrations and updates without having to bring applications down
    - Better configuration management
  - 3 environments (DMZ, Internal, Sandbox)
    - Sandbox environment allows admins to test upgrades and new capabilities before rolling out to production environments
  - Upgrade to Version 3.11
    - Stability and bug fixes
    - Ability to use highly available EMC Isilon storage
    - Security updates
    - Horizontal pod scaling
    - Grafana Dashboard
      - Working on making this available to application owners
    - Kubernetes Operators are supported

ASDC in the Cloud

- ESDIS heavily investing in commercial cloud
  - Primary motivations
    - To allow the science community to do unprecedented science with unconstrained processing and storage resources across all EOSDIS data in highly available AWS
    - The size of future datasets makes on-prem archiving highly undesirable
      - SWOT and NISAR most often cited
  - Building and AWS-native ingest and archive system called Cumulus

ASDC in the Cloud

- Phase 1 Cloud Proposal/Plan to ESDIS
  - ASDC developing a plan for a “Phase 1” cloud activity
    - Due to ESDIS this week
    - Covers 18 months beginning early May 2020
    - Concept for how to leverage cloud for supporting services (subsetting, etc)
    - Concept for ingesting data into the cloud
      - Ingest to ASDC, send a copy to AWS
    - Identifying some ASDC data sets to move into the cloud
      - Will include one or two “high value” data sets from CERES, MISR, and MOPITT
      - For CERES, we are using the SYN1deg for Terra-Aqua-MODIS
      - New missions are expected to be “born in the cloud”
        - MAIA and TEMPO
  - Data will be still be archived onsite at ASDC for the forseeable future
  - Does NOT include migrating CERES science data production

ASDC in the Cloud

- Guiding Principles
  - User experience will not be interrupted or degraded
  - Activities being planned in a way to avoid introducing technical and schedule risks and dependencies
  - ASDC will be working closely with ESDIS on identifying risks, technical complexities, and uncertainties, and proposing concepts and architectures for mitigating these concerns

ASDC in the Cloud

- We want to hear your thoughts!
  - What do you consider other “high value” ASDC data sets for access in AWS?
  - Value proposition for you and your colleagues?
  - Challenges with using the cloud?
  - Is AWS sufficient or do you have any preference for a different cloud vendor (i.e. Google or Microsoft)?

Questions?